

A Hamiltonian approach to sufficiency in optimal control with minimal regularity conditions: Part I

G. Stefani ¹ and P. Zezza ²

¹Dipartimento di Matematica e Informatica, Via di Santa Marta 3 - 50139 FIRENZE - ITALIA

²Dipartimento di Scienze per l'Economia e l'Industria, Via delle Pandette 9 - 50137 FIRENZE - ITALIA

July 15, 2015

Abstract

In this paper we develop a Hamiltonian approach to sufficient conditions in optimal control problems. We extend the known conditions for C^2 maximised Hamiltonians into two directions: on the one hand we explain the role of a super Hamiltonian (i.e. a Hamiltonian which is greater then or equal to the maximised one) on the other we develop the theory under some minimal regularity assumptions. The results we present enclose many known results and they can be used to tackle new problems.

Keywords. Sufficient condition, strong local minima, Hamiltonian methods.

MSC Classification. 49K15, 49J30

1 Introduction

This paper is the first part of a larger project whose goal is to describe how a Hamiltonian approach can be a key instrument in optimal control problems (OCPs).

Hamiltonian methods have been used in OCPs to state sufficient conditions ensuring the strong local optimality of a reference trajectory. The main feature of this approach is that it allows us to compare the costs of different trajectories by lifting them to the cotangent bundle and hence independently of the control values. The seminal idea goes back to 1879 when K. Weierstrass discovered a method which enables one to establish a strong minimum property for solutions of Euler-Lagrange's equations, i.e. for stationary curves. We can summarise Weierstrass's method as a combination of a local convexity assumption on the integrand and of an embedding of the given curve in a suitable field of non-intersecting stationary curves. The extension to optimal control requires some efforts since one has to deal with the maximised Hamiltonian coming from the Pontryagin Maximum Principle (PMP). The PMP was

introduced in 1956 by Lev Pontryagin and it is an extension to OCPs of the Euler-Lagrange's equation for the Calculus of Variations (CV). One of the main difficulty being that the maximised Hamiltonian in OCPs does not possess the same regularity properties as the one in the CV.

The goal of this project is to propose a unified Hamiltonian approach to sufficient optimality conditions for OCPs whose state evolves on a manifold M . These results include some known ones, emphasizing their common features, and will allow us to tackle new problems.

The leading ideas of the project are the following:

- To use the symplectic properties of the cotangent bundle to compare the costs of neighbouring admissible trajectories by lifting them to the cotangent bundle.
- To define in the cotangent bundle T^*M a suitable Hamiltonian flow \mathcal{H}_t emanating from a horizontal Lagrangian submanifold Λ . This flow is the one of the maximised Hamiltonian when this last is at least C^2 .
- To estimate the variation of the cost of admissible trajectories by the variation of a function of their final points and, if it is the case, their final times.
- To obtain a suitable second order approximation (*2nd variation*) in the form of a coordinate-free Linear-Quadratic (LQ) problem and to require its coercivity.
- To show that \mathcal{H}_{t*} (the derivative of \mathcal{H}_t along the reference extremal) is, up to an isomorphism, the linear Hamiltonian flow associated to the LQ problem.
- To use the coercivity of the *2nd variation* to substitute the manifold described by the transversality conditions, Λ_0 , by an horizontal one Λ . This can be obtained by adding a penalty term which reduces the problem to a problem with free initial point and whose *2nd variation* is still coercive, see M. Hestenes [16]. This allows us to overcome a difficult point: in optimal control problems, for lack of controllability it may happen that the projection of the flow starting from Λ_0 at $t = 0$ is not locally onto for a non-trivial time interval.
- To deduce for a problem with free initial point and fixed final point that the projection on M of \mathcal{H}_t emanating from Λ is locally invertible so that we can go back to the first issue and we can compare the costs of neighbouring admissible trajectories by lifting them to the cotangent bundle.
- To use again the coercivity of the *2nd variation* to complete the proof for the general case.

The project goes back to some initial papers [4] and [3] where the use of this kind of approach to OCPs for the case when the maximised Hamiltonian is at least C^2 did begin.

In [4] the authors studied a Bolza problem in \mathbb{R}^n on a fixed time interval and with general end-points constraints. They assumed that the data are uniformly quasi- C^2 , see Definition 1 therein, where the *quasi* refers to the fact that the data are time-dependent, moreover, under the strengthened Legendre condition, assuming that in a neighbourhood of the reference Pontryagin extremal the local maximum of

the Hamiltonian coincides with the global maximum they proved that the maximised Hamiltonian is quasi- C^2 .

This kind of regularity allows us to look for second order conditions which will guarantee that the abstract theory can be applied to obtain strong local optimality.

When the state evolves on manifolds, the first issue is to prove an invariant version of second order conditions; this problem has been addressed in [3] for a Mayer problem on a fixed time interval. This result can be obtained by pulling back the control problem to a neighbourhood of the initial point \hat{x}_0 , in this way the second order approximation leads naturally to derive a Hamiltonian formulation of the second variation as a linear-quadratic optimal control problem on $T_{\hat{x}_0}M$ with control functions in L^2 . This formulation is coordinate-free and hence invariant.

In the book [2], dedicated to the *geometric* approach to control problems, the smooth case is investigated in a Hamiltonian setting for a Bolza problem with fixed end points; in [9] the authors present a numerical algorithms to compute the first point where the trajectory ceases to be locally optimal.

The assumption that the data are uniformly quasi- C^2 can be weakened as shown in [5] by imposing conditions directly on the Hamiltonian flow (Theorem 2.3 therein), these conditions make sense even if the second variation does not exist and, when the second variation exists, they are implied by its coercivity.

Here we develop the first part of the project and we relax the regularity assumptions on the Hamiltonian in such a way that they guarantee the existence and the regularity of the flow so that this technique can be applied to a larger class of different OCPs. The Hamiltonian we consider can be either the maximised Hamiltonian or a suitable super-Hamiltonian which is not necessarily equal to the maximised one but which will be needed in the study of problems where the control contains a singular arc.

The regularity assumptions we propose allows us to use the symplectic properties of Hamiltonian flows (Lemma 3.3) to state abstract sufficient optimality conditions for a Bolza problem (Theorem 3.9) and we show their possible applications to a minimum time problem (Theorem 3.15).

In the final Section 4 we briefly summarise the optimal control problems where our assumptions are satisfied and this approach has been used.

In a forthcoming paper we will give a more detailed description of this approach and larger set of references, furthermore we will give some suggestions about possible extensions and further investigations. For different approaches here we quote only [8], [14], [19] and reference therein.

1.1 The problem

We consider the following optimisation problem

$$\begin{aligned} & \text{Minimise} \\ & J(T, \xi, v) = c_0(\xi(0)) + c_f(\xi(T)) + \int_0^T f^0(\xi(t), v(t)) dt \end{aligned} \quad (1.1)$$

subject to

$$\dot{\xi}(t) = f(\xi(t), v(t)), \quad \text{a.e. } t \in [0, T] \quad (1.2a)$$

$$\xi(t) \in M \quad (1.2b)$$

$$\xi(0) \in N_0, \quad \xi(T) \in N_f \quad (1.2c)$$

$$v \in L^\infty([0, T], U), \quad U \subseteq \mathbb{R}^m. \quad (1.2d)$$

The final time T can be fixed or variable, M is a n -dimensional connected paracompact smooth manifold M and the state end points constraints N_0, N_f are smooth connected embedded submanifolds of M . We assume that c_0, c_f, f^0, f are defined on open sets and that they are C^∞ .

We take smooth time independent data because we are interested in the irregularities arising from the maximisation of the Hamiltonian.

The couple (ξ, v) is called *admissible* if it is a solution of (1.2a) which satisfies the constraints (1.2b)–(1.2c)–(1.2d), in this case we refer to ξ as an *admissible trajectory* and to v as the *associated control*.

We consider as a *candidate optimal solution* a given reference admissible trajectory $\hat{\xi}$ which is identified by the triple $(\hat{T}, \hat{\xi}, \hat{v})$ and we study its strong local optimality according to the following Definition 1.1.

Definition 1.1. [*Strong local minimiser*]

The reference admissible trajectory $\hat{\xi}$ is a strong local minimiser of the above considered problem (1.1), if there are neighbourhoods $\mathcal{U} \subseteq \mathbb{R} \times M$ of the graph of $\hat{\xi}$, denoted by $\Gamma_{\hat{\xi}}$, and $\mathcal{O} \subseteq \mathbb{R} \times M$ of $(\hat{T}, \hat{\xi}(\hat{T}))$ such that $\hat{\xi}$ is a minimiser among all the admissible trajectories ξ satisfying

$$\Gamma_\xi \subset \mathcal{U}, \quad (T, \xi(T)) \in \mathcal{O}$$

independently of the associated control.

When the final time is fixed this definition reduces to the usual definition of strong local minimum which uses the C^0 topology to describe a neighbourhood of the admissible trajectories; whereas in the case of variable end time our definition is local with respect to the graph of $\hat{\xi}$, and hence local with respect to both the final time and the final point.

This notion has been called *time-state-local optimality* in [23]–[24], where it is used also a stronger version of optimality, called *state-local optimality*.

2 Notations and preliminary results

We recall that for every connected paracompact smooth manifold M (see [18]) there is a Riemannian structure which induces a distance d_M on M and the metric topology is the same as the original topology. This distance allows us to talk about Lipschitz property for maps on manifolds.

The next definition will be used to describe the main regularity assumption which is a strengthening of the usual Caratheodory-type assumption.

Definition 2.1. Assume that M, N are finite dimensional Riemannian smooth manifolds and J is an open interval in \mathbb{R} . We will say that the map $G : J \times M \rightarrow N$ is a Lipschitz–Carathéodory map if it satisfies the following

- i. For almost every $t \in J$ the map $x \mapsto G(t, x)$ is locally Lipschitz.
- ii. For each $x \in M$ the map $t \mapsto G(t, x)$ is bounded measurable.
- iii. For any compact set $K \subseteq M$ there is an essentially bounded measurable function m such that

$$d_N(G(t, x), G(t, y)) \leq m(t) d_M(x, y), \quad x, y \in K$$

When G is a time-dependent vector field and hence $N = TM$ is the tangent space to M , this hypothesis assures existence, uniqueness and Lipschitz continuity in (t, z_0) of the solutions of the differential equation

$$\dot{\zeta}(t) = G(t, \zeta), \quad \zeta(t_0) = z_0, \quad t_0 \in J.$$

Moreover we will use it in an analogous but simpler way when $N = \mathbb{R}$ to obtain that the Lipschitz continuity of $(t, z) \mapsto \int_{t_0}^t G(s, z) ds$.

Let f be a vector field on the manifold M and $\varphi : M \rightarrow \mathbb{R}$ be a smooth function. The action of f on φ (directional derivative or Lie derivative) evaluated at a point x is denoted with one the two expressions

$$L_f \varphi(x) = \langle d\varphi(x), f(x) \rangle.$$

For any C^2 -function φ such that $d\varphi(x) = 0$ the second derivative $D^2\varphi(x)$ is well defined as a bilinear symmetric form on $T_x M$.

Finally we identify any bilinear form Q on a vector space V with a linear form $Q : V \rightarrow V^*$ and we write

$$Q(v, w) := \langle Qv, w \rangle, \quad Q[v]^2 := Q(v, v).$$

Remark 2.2. We will sometimes omit writing explicitly that equalities between L^∞ functions hold almost everywhere unless we need to emphasise it.

2.1 Symplectic notations

For a general introduction to symplectic geometry and its application to variational problems we refer to [7] while for more specific applications to optimal control we refer to [1].

Denote by $\pi : T^*M \rightarrow M$ the cotangent bundle; it is well known that T^*M possesses a canonically defined symplectic structure, given by the symplectic form $\sigma_\ell = ds(\ell)$, where ℓ denotes an element of T^*M and s is the Liouville canonical 1-form $s(\ell) = \ell \circ \pi_*$.

If q_i are local coordinates on the base manifold M and p_i the fibre coordinates then in local coordinates we can write

$$s := \sum_{i=1}^n p_i dq_i, \quad \sigma := \sum_i dp_i \wedge dq_i,$$

where d denotes the exterior derivative and \wedge denotes the exterior product.

We recall that the symplectic structure allows us to associate, with each time dependent locally defined Hamiltonian $H_t : T^*M \rightarrow \mathbb{R}$ a unique vector field \vec{H}_t on T^*M defined by the action

$$\langle dH_t(\ell), \cdot \rangle = \sigma_\ell(\cdot, \vec{H}_t(\ell)).$$

This vector field defines a corresponding Hamiltonian system

$$\dot{\lambda}(t) = \vec{H}_t(\lambda(t)), \quad \lambda(0) = \ell, \quad a.e. \ t \in [0, T] \quad (2.1)$$

and we will denote its flow by $\ell \mapsto \mathcal{H}(t, \ell)$. Note that for any time dependent object we will use the notation

$$M(t, \ell) = M_t(\ell) = M_\ell(t)$$

when one of the variables is to be considered as fixed.

2.2 The Pontryagin maximum principle

We assume that the candidate optimal solution $\hat{\xi}$ is a state extremal and with this we mean that it satisfies the PMP, see Assumption 2.3.

For simpler notations we set

$$\hat{x}_0 := \hat{\xi}(0), \quad \hat{x}_f := \hat{\xi}(\hat{T})$$

To state the PMP we introduce three Hamiltonians; if \mathcal{U} is the common open domain of both f and f^0 let $\Omega := \pi^{-1}(\mathcal{U}) \subseteq T^*M$.

We define

- 1) the (control dependent) *pre-Hamiltonian* $F : \Omega \times U \rightarrow \mathbb{R}$ as

$$F : (\ell, u) \mapsto \langle \ell, f(\pi(\ell), u) \rangle - p_0 f^0(\pi(\ell), u),$$

- 2) the time dependent *reference Hamiltonian* $\hat{F} : [0, \hat{T}] \times \Omega \rightarrow \mathbb{R}$ as

$$\hat{F} : (t, \ell) \mapsto F(\ell, \hat{v}(t)),$$

- 3) the *maximised Hamiltonian* $F_{\max} : \Omega \rightarrow \overline{\mathbb{R}}$ as

$$F_{\max}(\ell) := \sup_{u \in U} F(\ell, u).$$

Let us remark that the maximised Hamiltonian could take the value $+\infty$.

All the above Hamiltonians depend on the parameter p_0 which can take the values $\{0, 1\}$ characterising, respectively, the abnormal and normal case.

Assumption 2.3. *[Pontryagin Maximum Principle] There is a non-trivial couple $(p_0, \hat{\lambda})$, where $p_0 \in \{0, 1\}$ and $\hat{\lambda} : [0, \hat{T}] \rightarrow T^*M$ is a lifting of $\hat{\xi}$ to T^*M , (i.e. $\pi \circ \hat{\lambda} = \hat{\xi}$) satisfying the Hamiltonian system*

$$\dot{\hat{\lambda}}(t) = \overrightarrow{\hat{F}_t}(\hat{\lambda}(t)), \quad \text{a.e. } t \in [0, \hat{T}], \quad (2.2)$$

the transversality conditions

$$\hat{\lambda}(0) = p_0 dc_0(\hat{x}_0) \quad \text{on } T_{\hat{x}_0}N_0 \quad (2.3)$$

$$\hat{\lambda}(\hat{T}) = -p_0 dc_f(\hat{x}_f) \quad \text{on } T_{\hat{x}_f}N_f, \quad (2.4)$$

and the maximisation property

$$\hat{F}_t(\hat{\lambda}(t)) = \max_{u \in U} F(\hat{\lambda}(t), u) = F_{\max}(\hat{\lambda}(t)).$$

Moreover $\hat{F}_t(\hat{\lambda}(t))$ is constant and it is zero when the end time T is variable.

$\hat{\lambda}$ is the Pontryagin extremal associated to the state extremal $\hat{\xi}$. For simpler notations we set

$$\hat{\ell}_0 := \hat{\lambda}(0), \quad \hat{\ell}_f := \hat{\lambda}(\hat{T}).$$

It is not difficult to see that the two functions $p_0 c_0$ and $p_0 c_f$ act on N_0 and N_f , respectively, but they can be extended to an open set in M in such a way that the transversality conditions (2.3) and (2.4) hold on the whole tangent space. Namely we denote by $\alpha, \beta : M \rightarrow \mathbb{R}$ two locally defined functions such that

$$\alpha = p_0 c_0 \text{ on } N_0, \quad \hat{\ell}_0 = d\alpha(\hat{x}_0) \quad (2.5)$$

$$\beta = p_0 c_f \text{ on } N_f, \quad \hat{\ell}_f = -d\beta(\hat{x}_f). \quad (2.6)$$

The transversality conditions can also be expressed as

$$\begin{aligned} \hat{\ell}_0 \in \Lambda_0 &:= \{d\alpha(q) + \omega : q \in N_0, \omega \in T_q^\perp N_0\} \\ \hat{\ell}_f \in \Lambda_f &:= \{-d\beta(q) + \omega : q \in N_f, \omega \in T_q^\perp N_f\} \end{aligned}$$

where " \perp " means orthogonal with respect to the dual coupling.

In the normal case α, β are cost functions equivalent to the original ones while in the abnormal case they are extensions of the zero function on the constraints.

When $p_0 = 0$ all the costs disappear from our conditions and indeed we will study a problem with a zero cost. Proving that $\hat{\xi}$ is a *strict* minimiser will imply that it is isolated among the admissible trajectories.

3 A Hamiltonian approach to optimality

In this section we extend the known results for a C^2 maximised Hamiltonian in two different directions. We first prove that the Cartan form maintains its symplectic properties under weaker regularity assumptions on the Hamiltonian, Section 3.1.

Afterwards, motivated by the singular case where the maximised Hamiltonian does not possess the required regularity properties, we introduce a super-Hamiltonian which can be used to compare the costs of admissible trajectories by lifting them to the cotangent bundle, Section 3.2. Using these results we prove abstract sufficient conditions, Section 3.3.

3.1 The Cartan form

Let $\alpha : M \rightarrow \mathbb{R}$ be a smooth function and let Λ be the graph of $d\alpha$ on a contractible open set. Λ is a Lagrangian submanifold of T^*M and it is known that

$$s|_{\Lambda} = d(\alpha \circ \pi).$$

Let \mathcal{O} be an open set in T^*M and $J \subseteq \mathbb{R}$ be an open interval, with $J \supset [0, \widehat{T}]$, and $\mathcal{O} \supset \Lambda$, we consider a time dependent Hamiltonian

$$H : J \times \mathcal{O} \subseteq \mathbb{R} \times T^*M \rightarrow \mathbb{R}.$$

Associated to this Hamiltonian we consider the Cartan form on $J \times \mathcal{O}$

$$\omega = s - H dt.$$

The following assumption describes a set of *minimal* regularity conditions which the Hamiltonian H has to guarantee to develop our approach. In specific examples the properties of H will be different but they will imply this kind of regularity.

Assumption 3.1. *Assume that*

1. *The flow*

$$(t, \ell) \in J \times \Lambda \mapsto \mathcal{H}(t, \ell) \in T^*M$$

is well defined and Lipschitz continuous.

2. *The function*

$$(t, \ell) \in J \times \Lambda \mapsto \langle \mathcal{H}_t(\ell), \pi_* \vec{H}_t \circ \mathcal{H}_t(\ell) \rangle - H_t \circ \mathcal{H}_t(\ell) \in \mathbb{R}$$

is Lipschitz Caratheodory (Definition 2.1)

Thanks to Assumption 3.1 the Cartan form defines a Lipschitz function $\theta : J \times \Lambda \rightarrow \mathbb{R}$ which will play a crucial role. Let

$$\begin{aligned} \theta(t, \ell) &:= (\alpha \circ \pi)(\ell) + \int_{\mathcal{H}_{\ell}|_{[0, t]}} \omega = \\ &= (\alpha \circ \pi)(\ell) + \int_0^t \langle \mathcal{H}_s(\ell), \pi_* \vec{H}_s \circ \mathcal{H}_s(\ell) \rangle - H_s \circ \mathcal{H}_s(\ell) ds \end{aligned} \quad (3.1)$$

and hence

$$\partial_t \theta(t, \ell) = \langle \mathcal{H}_t(\ell), \pi_* \vec{H}_t \circ \mathcal{H}_t(\ell) \rangle - H_t \circ \mathcal{H}_t(\ell) \quad a.e \ t \in J \quad (3.2)$$

Let us now prove that, also in this case, the classical property of the Cartan form holds, that is the form $\mathcal{H}^*\omega$ is exact on $J \times \Lambda$.

We will consider differential 1-forms with possibly L^∞ coefficients; these forms are called (Whitney) flat forms and their properties are presented in [15]. The general theory is fully presented in Whitney's monograph [28], see also [13]. We do not use the results in their full extent, what we really need is the chain rule for Lipschitz functions as it can be found in Lemma 4.6.3 in [17] and a suitable version of the Stokes Theorem; an explicit proof is in [4] but the essential elements needed for the proof are contained in the cited books, [13] for the general case or [15] for the Lipschitz one.

The following lemma is the first step to prove that the form $\mathcal{H}^*\omega$ is exact on $J \times \Lambda$.

Lemma 3.2. *For every given $t \in J$ the form \mathcal{H}_t^*s is exact on Λ and*

$$\mathcal{H}_t^*s(\ell) = d\theta_t(\ell) \quad \text{a.e. } \ell \in \Lambda \quad (3.3)$$

Proof. It is equivalent to prove that for every Lipschitz curve $\gamma : [0, 1] \rightarrow \Lambda$

$$\int_\gamma \mathcal{H}_t^*s = \int_\gamma d\theta_t = \theta_t(\gamma(1)) - \theta_t(\gamma(0)).$$

Let us consider the set

$$\Delta := \left\{ (\tau, s) : 0 \leq \tau \leq t, 0 \leq s \leq 1 \right\}$$

and the map

$$\phi : (\tau, s) \mapsto \mathcal{H}(\tau, \gamma(s)).$$

by the Stokes theorem we have

$$\begin{aligned} \int_\Delta \phi^* \sigma &= \int_{\phi|_{s=0}} s + \int_{\phi|_{\tau=t}} s - \int_{\phi|_{s=1}} s - \int_{\phi|_{\tau=0}} s = \\ &= \int_{\mathcal{H}_{\gamma(0)}|_{[0, t]}} s + \int_{\mathcal{H}_t \circ \gamma|_{[0, 1]}} s - \int_{\mathcal{H}_{\gamma(1)}|_{[0, t]}} s - \int_{\gamma|_{[0, 1]}} s \end{aligned}$$

moreover

$$\begin{aligned} \int_\Delta \phi^* \sigma &= \int_0^t \left\{ \int_0^1 \sigma \left(\vec{H}_\tau \circ \mathcal{H}_\tau \circ \gamma(s), \partial_s \phi(\tau, s) \right) ds \right\} d\tau = \\ &= - \int_0^t \left\{ \int_0^1 \langle dH_\tau \circ \phi(\tau, s), \partial_s \phi(\tau, s) \rangle ds \right\} d\tau = \\ &= - \int_0^t H_\tau \circ \mathcal{H}_{\gamma(1)} d\tau + \int_0^t H_\tau \circ \mathcal{H}_{\gamma(0)} d\tau \end{aligned}$$

by equating the two right hand sides we obtain

$$\begin{aligned} \int_{\mathcal{H}_t \circ \gamma|_{[0, 1]}} s &= \int_{\gamma|_{[0, 1]}} s + \int_{\mathcal{H}_{\gamma(1)}|_{[0, t]}} \omega - \int_{\mathcal{H}_{\gamma(0)}|_{[0, t]}} \omega = \\ &= \theta(t, \gamma(1)) - \theta(t, \gamma(0)). \end{aligned}$$

□

We are now able to prove the main result of this section

Lemma 3.3. *The differential form $\mathcal{H}^*\omega$ is exact on $J \times \Lambda$ and*

$$\mathcal{H}^*\omega = d\theta \quad \text{a.e. } (t, \ell) \in J \times \Lambda$$

Proof. We can equivalently prove that for every Lipschitz curve

$$\mu : t \in [0, 1] \mapsto (t, \gamma(t)) \in J \times \Lambda$$

we have

$$\int_{\mu} \mathcal{H}^*\omega = \int_{\mu} d\theta = \theta(1, \gamma(1)) - \theta(0, \gamma(0)).$$

By definition we have that

$$\begin{aligned} \int_{\mu} \mathcal{H}^*\omega &= \int_{\mu} \mathcal{H}^*s - \int_0^1 H(s, \mathcal{H}_s \circ \gamma(s)) ds = \\ &= \int_0^1 \langle \mathcal{H}_s \circ \gamma(s), \frac{d}{ds} \pi \mathcal{H}_s \circ \gamma(s) \rangle ds - \int_0^1 H(s, \mathcal{H}_s \circ \gamma(s)) ds = \\ &= \int_0^1 \left\{ \langle \mathcal{H}_s \circ \gamma(s), \pi_* \vec{H}_s \circ \mathcal{H}_s \circ \gamma(s) \rangle - H(s, \mathcal{H}_s \circ \gamma(s)) \right\} ds + \\ &+ \int_0^1 \langle \mathcal{H}_s \circ \gamma(s), \pi_* \mathcal{H}_{s*} \dot{\gamma}(s) \rangle ds \end{aligned}$$

and by (3.2) and (3.3)

$$\begin{aligned} \int_{\mu} \mathcal{H}^*\omega &= \int_{\mu} \partial_s \theta(s, \ell) ds + \int_{\mu} d\theta_s(\ell) = \\ &= \int_{\mu} d\theta(s, \ell) = \theta(1, \gamma(1)) - \theta(0, \gamma(0)). \end{aligned}$$

□

When the state projection of the Hamiltonian flow is invertible we obtain some important properties of the function θ . We recall that a homeomorphism between metric spaces is said to be bi-Lipschitz if it is Lipschitz and has a Lipschitz inverse.

Lemma 3.4. *Suppose that, for a given $t \in J$, the function $\pi \mathcal{H}_t$ is bi-Lipschitz then on $\pi \mathcal{H}_t(\Lambda)$*

$$(i) \quad d\left(\theta_t \circ (\pi \mathcal{H}_t)^{-1}\right) = \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}.$$

(ii) *The function $\alpha_t := \theta_t \circ (\pi \mathcal{H}_t)^{-1}$ is a C^1 function.*

(iii) $\Lambda_t := \mathcal{H}_t(\Lambda)$ *is the graph of $d\alpha_t$.*

Proof.

Since $\theta_t \circ (\pi \mathcal{H}_t)^{-1}$ is Lipschitz then by the chain rule we can prove (i) by proving

that their integrals coincide over any Lipschitz curve $\gamma : [0, 1] \rightarrow \pi\mathcal{H}_t(\Lambda)$.
By (3.3) we have

$$\begin{aligned}
\int_{\gamma} d\left(\theta_t \circ (\pi\mathcal{H}_t)^{-1}\right) &= \int_0^1 \langle d\left(\theta_t \circ (\pi\mathcal{H}_t)^{-1}\right)(\gamma(s)), \dot{\gamma}(s) \rangle ds = \\
&= \int_0^1 \langle d\theta_t \circ (\pi\mathcal{H}_t)^{-1}(\gamma(s)), (\pi\mathcal{H}_t)_*^{-1} \dot{\gamma}(s) \rangle ds = \\
&= \int_{(\pi\mathcal{H}_t)^{-1} \circ \gamma} d\theta_t = \int_{(\pi\mathcal{H}_t)^{-1} \circ \gamma} \mathcal{H}_t^* s = \\
&= \int_0^1 \langle \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}(\gamma(s)), \frac{d}{ds}(\pi\mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1} \circ \gamma(s)) \rangle ds = \\
&= \int_0^1 \langle \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}(\gamma(s)), \dot{\gamma}(s) \rangle ds
\end{aligned}$$

which proves statement (i). (ii) Follows immediately from (i). To prove (iii) it is sufficient to notice that

$$\Lambda_t = \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}(\pi\mathcal{H}_t(\Lambda)).$$

□

3.2 The super Hamiltonian and its properties.

Throughout this section we assume that α satisfies (2.5) and that Λ is the graph of $d\alpha$ on a contractible neighbourhood of \hat{x}_0 . Moreover we assume that H satisfies Assumption 3.1 and the following Assumption 3.5, this last motivates the name of super-Hamiltonian. We underline that these assumptions concern jointly the super-Hamiltonian and the horizontal Lagrangian manifold Λ .

Assumption 3.5. *The Hamiltonian H satisfies the following*

$$H_t \circ \mathcal{H}_t(\ell) \geq F_{\max}(\mathcal{H}_t(\ell)), \quad \ell \in \Lambda \quad (3.4a)$$

$$H_t \circ \hat{\lambda}(t) = \hat{F}_t \circ \hat{\lambda}(t) = F_{\max} \circ \hat{\lambda}(t), \quad a.e. \ t \in J \quad (3.4b)$$

$$\vec{H}_t(\hat{\lambda}(t)) = \vec{\hat{F}}_t(\hat{\lambda}(t)), \quad a.e. \ t \in J. \quad (3.4c)$$

As a consequence the Pontryagin extremal $\hat{\lambda}$ is also a solution of the system

$$\dot{\lambda}(t) = \vec{H}_t(\lambda(t)) \quad a.e. \ t \in [0, \hat{T}]. \quad (3.5)$$

Remark 3.6. If the maximised Hamiltonian F_{\max} satisfies Assumption 3.1 then it satisfies also Assumption 3.5; for example in the classical case when F_{\max} is C^2 we can take $H := F_{\max}$ for any Λ .

From Assumption 3.1 it follows that there exists an interval $I := [0, T]$ with $T > \hat{T}$ such that $\mathcal{H}_t(\hat{\ell}_0)$ is defined for $t \in I$ and we have $\mathcal{H}_t(\hat{\ell}_0) = \hat{\lambda}(t)$ on $[0, \hat{T}]$. Moreover from the compactness of the time interval I , it follows that there is an

open neighbourhood $\mathcal{O}_{\widehat{\ell}_0}$ of $\widehat{\ell}_0$ such that, without loss of generality, we can redefine $\Lambda := \Lambda \cap \mathcal{O}_{\widehat{\ell}_0}$, to obtain that \mathcal{H} is defined on $I \times \Lambda$ and $\widehat{\mathcal{F}}$ is defined on $[0, \widehat{T}] \times \Lambda$, where $\widehat{\mathcal{F}}$ is the flow of \widehat{F}_t .

To prove the main Theorem we need a final crucial assumption concerning the invertibility of the projection of the flow of the Hamiltonian onto the state space.

Assumption 3.7. *Assume that the function*

$$id_I \times \pi\mathcal{H} : (t, \ell) \in I \times \Lambda \mapsto (t, \pi\mathcal{H}_t(\ell)) \in I \times M$$

is bi-Lipschitz between $I \times \Lambda$ and an open set \mathcal{U} of $I \times M$ containing the graph of $\widehat{\xi}$.

To verify the bi-Lipschitz assumption one can use one of the available inverse function theorems for locally Lipschitz functions which are based on the local invertibility properties of $\pi_*\mathcal{H}_{t*}$ in $[0, \widehat{T}] \times \{\widehat{\ell}_0\}$. We refer, for example, to the one proposed by F. Clarke in [12].

Remark 3.8. For optimal control problems where the initial point is free the problem is always normal, moreover the initial Lagrangian manifold Λ_0 is itself horizontal and we take it as Λ . In general, however, this is not the case and one has to define an appropriate α , in our approach we expect that the suitable α can be obtained from the second order conditions as it is the case in the applications we describe in the final Section 4.

We can now state the main Theorem which compares the reference cost with the cost of an admissible neighbouring trajectory.

Theorem 3.9. *Under Assumption 3.1–3.5–3.7, let $\xi : [0, T] \rightarrow M$ be an admissible trajectory whose graph is contained in \mathcal{U} and let $\rho(t) := (\pi\mathcal{H}_t)^{-1} \circ \xi(t)$ then*

$$p_0 \left(J(T, \xi, v) - J(\widehat{T}, \widehat{\xi}, \widehat{v}) \right) \geq \beta(\xi(T)) + \theta(T, \rho(T)) - \beta(\widehat{x}_f) - \theta(\widehat{T}, \widehat{\ell}_0) \quad (3.6)$$

Proof. Let $\mu := \mathcal{H}_t \circ \rho$ be the lift of ξ to T^*M , by Lemma 3.3 we obtain

$$\begin{aligned} 0 &= \int_0^T \left\{ \langle \mu(t), \dot{\xi}(t) \rangle - H_t \circ \mu(t) \right\} dt + \theta(\widehat{T}, \widehat{\ell}_0) - \theta(T, \rho(T)) + \\ &\quad - \int_0^{\widehat{T}} \left\{ \langle \widehat{\lambda}(t), \widehat{f}_t \circ \widehat{\xi}(t) \rangle - \widehat{F}_t \circ \widehat{\lambda}(t) \right\} dt + \alpha(\pi(\mu(0))) - \alpha(\pi(\widehat{\ell}_0)) = \\ &= \int_0^T \left\{ \langle \mu(t), \dot{\xi}(t) \rangle - H_t \circ \mu(t) \right\} dt + \theta(\widehat{T}, \widehat{\ell}_0) - \theta(T, \rho(T)) + \\ &\quad + \int_0^T p_0 f^0(\xi(t), v(t)) dt - \int_0^T p_0 f^0(\xi(t), v(t)) dt + \\ &\quad - \int_0^{\widehat{T}} \left\{ \langle \widehat{\lambda}(t), \widehat{f}_t \circ \widehat{\xi}(t) \rangle - \widehat{F}_t \circ \widehat{\lambda}(t) \right\} dt + \alpha(\xi(0)) - \alpha(\widehat{x}_0) + \\ &\quad + \int_0^{\widehat{T}} p_0 f^0(\widehat{\xi}(t), \widehat{v}(t)) dt - \int_0^{\widehat{T}} p_0 f^0(\widehat{\xi}(t), \widehat{v}(t)) dt. \end{aligned}$$

If we isolate, on the left hand side, the terms which describe the costs of the two trajectories we obtain

$$\begin{aligned}
& \alpha(\xi(0)) + \int_0^T p_0 f^0(\xi(t), v(t)) dt - \alpha(\widehat{\xi}(0)) - \int_0^{\widehat{T}} p_0 f^0(\widehat{\xi}(t), \widehat{v}(t)) dt = \\
& = - \int_0^T \{ \langle \mu(t), f(\xi(t), v(t)) \rangle - p_0 f^0(\xi(t), v(t)) - H_t \circ \mu(t) \} dt + \\
& + \int_0^{\widehat{T}} \{ \langle \widehat{\lambda}(t), f(\widehat{\xi}(t), \widehat{v}(t)) \rangle - p_0 f^0(\widehat{\xi}(t), \widehat{v}(t)) - \widehat{F}_t \circ \widehat{\lambda}(t) \} dt + \\
& - \theta(\widehat{T}, \widehat{\ell}_0) + \theta(T, \rho(T)) = \\
& = - \int_0^T \{ F(\mu(t), v(t)) - H_t \circ \mu(t) \} dt - \theta(\widehat{T}, \widehat{\ell}_0) + \theta(T, \rho(T))
\end{aligned}$$

where we have used the definition of F and of \widehat{F} . Now by the (3.4a) property of the super Hamiltonian, by adding to both sides $\beta(\xi(T)) - \beta(\widehat{x}_f)$, and recalling that on the initial manifold $\alpha = p_0 c_0$ and on the final manifold $\beta = p_0 c_f$ we obtain

$$p_0 \left(J(T, \xi, v) - J(\widehat{T}, \widehat{\xi}, \widehat{v}) \right) \geq \beta(\xi(T)) + \theta(T, \rho(T)) - \beta(\widehat{x}_f) - \theta(\widehat{T}, \widehat{\ell}_0)$$

□

3.3 Abstract sufficient optimality conditions.

We are now able to state an abstract theorem which reduces the strong local optimality of an admissible reference trajectory $\widehat{\xi}$ to the local optimality of a suitable function of the right end point $(\widehat{T}, \widehat{x}_f)$ of the graph of $\widehat{\xi}$.

Let

$$\Phi: (t, x) \in \mathcal{U} \mapsto \beta(x) + \theta(t, (\pi \mathcal{H}_t)^{-1}(x)) = (\beta + \alpha_t)(x) \in \mathbb{R},$$

where $\alpha_t = \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}$ was defined in part (ii) of Lemma 3.4. By means of this function we can rewrite the inequality (3.6) as

$$p_0 \left(J(T, \xi, v) - J(\widehat{T}, \widehat{\xi}, \widehat{v}) \right) \geq \Phi(T, \xi(T)) - \Phi(\widehat{T}, \widehat{x}_f). \quad (3.7)$$

This relation will allow us to characterise the minima of the function $p_0 J$ by studying the function Φ at the reference point $(\widehat{T}, \widehat{x}_f)$. Indeed we can now state a sufficient optimality condition

Theorem 3.10. *In the normal case, $p_0 = 1$, under Assumptions 3.1–3.5–3.7, if there exists a neighbourhood \mathcal{O} of $(\widehat{T}, \widehat{x}_f)$ in N_f such that $(\widehat{T}, \widehat{x}_f)$ is a minimiser of Φ restricted to \mathcal{O} then $(\widehat{T}, \widehat{\xi}, \widehat{v})$ is a strong local minimiser.*

Proof. It follows immediately from (3.7). □

Remark 3.11. When the initial point is free we have that $\alpha = c_0$, if, moreover, the final point and time are fixed then the right hand side of (3.7) is zero and we can prove that $\widehat{\xi}$ is a strong local minimiser by only verifying Assumptions 3.1–3.5–3.7 for the given Λ . In the other cases we have to find a suitable Λ and to prove the minimality property of Φ at $(\widehat{T}, \widehat{x}_f)$.

Remark 3.12. Concerning the abnormal case, $p_0 = 0$, if one can prove that the minimum is strict then, as a byproduct, it follows that $\widehat{\xi}$ is isolated among the admissible trajectories. To prove that $\widehat{\xi}$ is a strict strong local minimiser the first step is to require that $(\widehat{T}, \widehat{x}_f)$ is a strict local minimiser for Φ , in this case, if we have another minimiser (T, ξ, v) , we can conclude that $T = \widehat{T}$ and $\xi(T) = \widehat{x}_f$. From the proof of Theorem 3.9 we have

$$0 = p_0 \left(J(T, \xi, v) - J(\widehat{T}, \widehat{\xi}, \widehat{v}) \right) = - \int_0^{\widehat{T}} \{ F(\mu(t), v(t)) - H_t \circ \mu(t) \} dt.$$

By Assumption (3.4a) we obtain that

$$F(\mu(t), v(t)) - H_t \circ \mu(t) = 0, \quad a.e. \ t \in [0, \widehat{T}].$$

On the other hand, since $\mu = \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1} \circ \xi$ we have that

$$\dot{\mu}(t) = \vec{H}_t(\mu(t)) - \mathcal{H}_{t*}(\pi \mathcal{H}_t)^{-1} \left(\pi_* \vec{H}_t(\mu(t)) - f(\xi(t), \mu(t)) \right) \quad a.e. \ t \in [0, \widehat{T}].$$

If

$$\pi_* \vec{H}_t(\mu(t)) - f(\xi(t), \mu(t)) = 0 \quad a.e. \ t \in [0, \widehat{T}] \quad (3.8)$$

then one can conclude that $\mu = \widehat{\lambda}$ since they both satisfy the same Hamiltonian equation with the same boundary conditions. Unfortunately Assumption (3.4a) is too mild to prove (3.8). We can strengthen it by assuming that

$$H_t(\ell) \geq F_{\max}(\ell), \quad \ell \in \mathcal{O} \quad (3.9)$$

where \mathcal{O} is a neighbourhood of the range of $\widehat{\lambda}$; we note that this is true if $H_t = F_{\max}$. From this new assumption we deduce that

$$F(\ell, v) - H_t(\ell) = 0, \quad \forall \ell : \pi \ell = \xi(t), \quad a.e. \ t \in [0, \widehat{T}], \quad (3.10)$$

hence by (3.10)

$$0 = \partial_v F(\mu(t), v) - \partial_v H_t(\mu(t)) = \pi_* \vec{H}_t(\mu(t)) - f(\xi(t), \mu(t)), \quad (3.11)$$

where ∂_v is the vertical derivative along the fibre. This is not the only way to obtain local uniqueness of the strong minimiser; for the case of singular control see for example [24], [10].

To state necessary and/or sufficient condition for $(\widehat{T}, \widehat{x}_f)$ to be a local minimiser for Φ we compute its first and second derivatives. We note that, by Lemma 3.4, the function Φ is a C^1 function of x at t fixed and Lemma 3.13 states that the point $(\widehat{T}, \widehat{x}_f)$ is a critical point for the function Φ without further assumptions on the data. On the other hand to obtain the existence of second derivatives we will require stronger regularity assumptions on the data which here we do not specify. We underline that since the first derivatives are zero, then the second ones are well defined as a quadratic function on $\mathbb{R} \times T_{\widehat{x}_f} M$

Lemma 3.13. *Under Assumptions 3.1–3.5–3.7, we have that*

- (i) $\partial_x \Phi(t, x) = d\beta(x) + \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x) = d(\beta + \alpha_t)(x), \quad t \in I$
- (ii) $\partial_t \Phi(t, x) = -H_t \circ \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x) = -H_t \circ (d\alpha_t)(x), \quad \text{a.e. } t \in I.$

Moreover

$$d\Phi(\hat{T}, \hat{x}_f) = 0.$$

Proof. (i) follows immediately from Lemma (3.4). To prove (ii) from (3.2) and Lemma (3.4) we have

$$\begin{aligned} \partial_t \Phi(t, x) &= \partial_t \theta(t, (\pi \mathcal{H}_t)^{-1}(x)) - \langle d\left(\theta_t \circ (\pi \mathcal{H}_t)^{-1}\right)(x), \pi_* \vec{H}_t \circ \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x) \rangle = \\ &= \partial_t \theta(t, (\pi \mathcal{H}_t)^{-1}(x)) - \langle \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x), \pi_* \vec{H}_t \circ \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x) \rangle = \\ &= -H_t \circ \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(x). \end{aligned}$$

Computing these derivatives at (\hat{T}, \hat{x}_f) , by the transversality condition (2.4), we obtain

$$\begin{aligned} \partial_x \Phi(\hat{T}, \hat{x}_f) &= d\beta(\hat{x}_f) + \mathcal{H}_{\hat{T}} \circ (\pi \mathcal{H}_{\hat{T}})^{-1}(\hat{x}_f) = \\ &= d\beta(\hat{x}_f) + \hat{\ell}_f = 0. \end{aligned}$$

By Assumption (3.4b) and by the PMP with variable final time we have

$$\partial_t \Phi(\hat{T}, \hat{x}_f) = -H_{\hat{T}}(\hat{\ell}_f) = -\hat{F}_{\hat{T}}(\hat{\ell}_f) = 0.$$

When the final time is fixed we are not interested in this derivative. \square

Lemma 3.14. *Let Assumptions 3.1–3.5–3.7 hold true and assume moreover that Φ is C^2 in a neighbourhood \mathcal{O} of (\hat{T}, \hat{x}_f) , then*

1. $\partial_{xx} \Phi(\hat{T}, \hat{x}_f) [\delta x]^2 = D^2(\beta + \alpha_{\hat{T}})(\hat{x}_f) [\delta x]^2 = \sigma((d\alpha_{\hat{T}})_* \delta x, d(-\beta)_* \delta x)$
2. $\partial_{tx} \Phi(\hat{T}, \hat{x}_f) \delta x = -L_{\delta x} L_{\hat{f}_{\hat{T}}} \alpha_{\hat{T}}(\hat{x}_f) = \sigma\left(\vec{H}_{\hat{T}}(\hat{\ell}_f), (d\alpha_{\hat{T}})_* \delta x\right)$
3. $\begin{aligned} \partial_{tt} \Phi(\hat{T}, \hat{x}_f) &= -\partial_t H(\hat{T}, \hat{x}_f) - L_{\hat{f}_{\hat{T}}}^2 \alpha_{\hat{T}}(\hat{x}_f) = \\ &= -\left(\partial_t H(\hat{T}, \hat{x}_f) + \sigma\left((d\alpha_{\hat{T}})_* \hat{f}_{\hat{T}}, \vec{H}_{\hat{T}}(\hat{\ell}_f)\right)\right). \end{aligned}$

Notice that since $\alpha_{\hat{T}} = \mathcal{H}_{\hat{T}} \circ (\pi \mathcal{H}_{\hat{T}})^{-1}$ then $(d\alpha_{\hat{T}})_* = \mathcal{H}_{\hat{T}*} (\pi \mathcal{H}_{\hat{T}})^{-1}_*$.

Proof.

1. It follows directly from Lemma 3.13, part (i).

2.-3. From Lemma 3.13, part (ii), taking into account that $\vec{H}_{\hat{T}}(\hat{\ell}_f) = \vec{F}_{\hat{T}}(\hat{\ell}_f)$ and some known symplectic equalities, that can be easily obtained in coordinates, we get

$$\begin{aligned} \partial_{tx} \Phi(\hat{T}, \hat{x}_f) \delta x &= -\sigma\left((d\alpha_{\hat{T}})_* \delta x, \vec{H}_{\hat{T}}(\hat{\ell}_f)\right) = \\ &= -L_{\delta x} L_{\hat{f}_{\hat{T}}} \alpha_{\hat{T}}(\hat{x}_f) \end{aligned}$$

and

$$\begin{aligned}
\partial_{tt}\Phi(\widehat{T}, \widehat{x}_f) &= \partial_t \left(-H(t, \mathcal{H}_t(\pi\mathcal{H}_t)^{-1}(\widehat{x}_f)) \right) = \\
&= -\partial_t H(\widehat{T}, \widehat{x}_f) + \sigma \left(\mathcal{H}_{\widehat{T}*}(\pi\mathcal{H}_{\widehat{T}})^{-1} \pi_* \vec{H}_{\widehat{T}}(\widehat{\ell}_f), \vec{H}_{\widehat{T}}(\widehat{\ell}_f) \right) = \\
&= -\partial_t H(\widehat{T}, \widehat{x}_f) + \sigma \left((d\alpha_{\widehat{T}})_* \widehat{f}_{\widehat{T}}(\widehat{x}_f), \vec{H}_{\widehat{T}}(\widehat{\ell}_f) \right) = \\
&= -\partial_t H(\widehat{T}, \widehat{x}_f) + L_{\widehat{f}_{\widehat{T}}}^2 \alpha_{\widehat{T}}(\widehat{x}_f).
\end{aligned}$$

□

3.4 The Minimum Time Problem

In this section we apply the results of the previous sections to the minimum time problem which is a special case which can be obtained from the general one by setting

$$c_0 = c_f = 0, \quad f^0 = 1.$$

For the minimum time problem we obtain a sufficient condition which do not involve the second derivatives of Φ , but requires more regularity of the super-Hamiltonian "near the final time" and a second assumption which is clearly verified when final point is fixed.

Theorem 3.15. *Let Assumptions 3.1–3.5–3.7 hold true and assume moreover that*

1. *The function $(t, x) \mapsto H_t \circ (\pi\mathcal{H}_t)^{-1}(x)$ is Lipschitz in a neighbourhood of $(\widehat{T}, \widehat{x}_f)$ in $\mathbb{R} \times M$, with a Lipschitz constant given by $L > 0$.*
2. *$\mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}(x) \in T_x^\perp N_f$, $(t, x) \in \mathcal{O}_{(\widehat{T}, \widehat{x}_f)}$, $x \in N_f$*

if $p_0 = 1$ (normal case), then $(\widehat{T}, \widehat{\xi})$ is a strong local minimum for the minimum time problem.

Proof. Let k be the dimension of the submanifold N_f , we can take local coordinates centered at \widehat{x}_f such that N_f is homomorphic to the plane given by the last $n - k$ coordinates equal to zero, so that we can consider the problem on $\mathbb{R} \times \mathbb{R}^n$ with $\widehat{x}_f = 0$ and $\xi(T) \in \mathbb{R}^k$.

For the minimum time problem in the normal case equation (3.7) reads

$$T - \widehat{T} \geq \Phi(T, \xi(T)) - \Phi(\widehat{T}, \widehat{x}_f) = \theta(T, (\pi\mathcal{H}_t)^{-1}(\xi(T))) - \theta(\widehat{T}, \widehat{x}_f). \quad (3.12)$$

By contradiction assume there exist admissible trajectories $\xi_n : [0, t_n] \rightarrow M$, $t_n < \widehat{T}$, whose graph is contained in \mathcal{U} and such that

$$\widehat{T} - t_n \rightarrow 0, \quad \text{and} \quad \|\xi_n(t_n)\| \rightarrow 0. \quad (3.13)$$

Consider the curve

$$\gamma := t \in [t_n, \widehat{T}] \mapsto \frac{\widehat{T} - t}{\widehat{T} - t_n} \xi_n(t_n)$$

which is such that $\gamma(t_n) = \xi_n(t_n)$ and $\gamma(\hat{T}) = 0$. By Lemma 3.3 and Assumption (ii) we get from equation (3.12)

$$\begin{aligned} t_n - \hat{T} &\geq \int_{t_n}^{\hat{T}} \left\{ \langle \mathcal{H}_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)), \dot{\gamma}(t) \rangle - H_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)) \right\} dt = \\ &= - \int_{t_n}^{\hat{T}} H_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)) dt \end{aligned}$$

Since $(\pi \mathcal{H}_{\hat{T}})^{-1}(\hat{x}_f) = \hat{\ell}_0$ by (3.4b) and Assumption (i) we can write

$$\begin{aligned} \hat{T} - t_n &\leq \int_{t_n}^{\hat{T}} H_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)) dt = \\ &= \int_{t_n}^{\hat{T}} \left(H_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)) - H_{\hat{T}} \circ (\pi \mathcal{H}_{\hat{T}})^{-1}(\hat{x}_f) \right) dt \leq \\ &\leq \int_{t_n}^{\hat{T}} \left\| H_t \circ (\pi \mathcal{H}_t)^{-1}(\gamma(t)) - H_{\hat{T}} \circ (\pi \mathcal{H}_{\hat{T}})^{-1}(\hat{x}_f) \right\| dt \leq \\ &\leq \int_{t_n}^{\hat{T}} L \left(\hat{T} - t + \frac{\hat{T} - t}{\hat{T} - t_n} \|\xi_n(t_n)\| \right) dt = \\ &= L \frac{(\hat{T} - t_n)^2}{2} \left(1 + \frac{\|\xi_n(t_n)\|}{\hat{T} - t_n} \right). \end{aligned}$$

Dividing by $\hat{T} - t_n > 0$ we obtain

$$1 \leq \frac{L}{2} (\hat{T} - t_n + \|\xi_n(t_n)\|)$$

which yield a contradiction by (3.13). \square

4 Final Comments

This general unified Hamiltonian approach can be used in different situations which has been already addressed in some published papers, where second order conditions have also been developed and appropriate H and Λ have been defined to obtain sufficient conditions for strong local optimality, here we briefly summarise them.

- **Bang–bang control.**

In this case $H := F_{\max}$ is continuous and \mathcal{H} piecewise C^∞ and Lipschitz. The case when there are a finite number of simple switches is studied in [6] for a Mayer problem with variable end–points on a fixed time interval and in [20] for a Bolza problem, while the corresponding minimum time problem is studied in [23]. The double switch case is addressed in [21] for a Mayer problem and in [22] for the minimum time problem. The numerical analysis with Maple of a case study is in [27].

- **Totally singular control**

In [26], for a Mayer problem and a single-input system the author first introduced a super-Hamiltonian, this step was essential because the maximised Hamiltonian does not define a regular flow. In [24] and [25] the result for the minimum time problem has been obtained as a by-product of problems with controls containing bang and singular arcs. Furthermore in [11] the minimum time problem has been studied in the multi-input case for a system where the controlled vector fields generate an involutive Lie algebra, while in [10] this last assumption is removed. We notice that the super Hamiltonian, and hence its flow, is C^∞ in a neighbourhood of the range of $\hat{\lambda}$ but it satisfies Assumption 3.5 only starting from a submanifold of T^*M containing Λ , as we require.

- **Bang-singular control**

The last application we consider is the case when the reference trajectory contains both singular and bang arcs. The obtained results concern the minimum time problem for a single-input control system. In [24] a bang-singular-bang trajectory in a problem with fixed end points is addressed, while in [25] it is considered a bang-singular trajectory with the initial point fixed and the final one constrained to the integral line of the controlled vector field. In this case the flow of the super Hamiltonian is sufficiently regular and satisfies Assumption 3.5 only starting from Λ .

References

- [1] A. A. Agrachev and R. V. Gamkrelidze. The exponential representation of flows and the chronological calculus. *Mat. Sb. (N.S.)*, 107(149)(4):467–532, 639, 1978.
- [2] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.
- [3] A. A. Agrachev, G. Stefani, and P. Zezza. An invariant second variation in optimal control. *Internat. J. Control*, 71(5):689–715, 1998.
- [4] A. A. Agrachev, G. Stefani, and P. Zezza. Strong minima in optimal control. *Tr. Mat. Inst. Steklova*, 220(Optim. Upr., Differ. Uravn. i Gladk. Optim.):8–26, 1998.
- [5] A. A. Agrachev, G. Stefani, and P. Zezza. A Hamiltonian approach to strong minima in optimal control. In *Differential geometry and control (Boulder, CO, 1997)*, volume 64 of *Proc. Sympos. Pure Math.*, pages 11–22. Amer. Math. Soc., Providence, RI, 1999.
- [6] A. A. Agrachev, G. Stefani, and P. Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optim.*, 41(4):991–1014 (electronic), 2002.
- [7] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical aspects of classical and celestial mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, third edition, 2006. [Dynamical systems. III], Translated from the Russian original by E. Khukhro.

- [8] M. S. Aronna, J. F. Bonnans, A. V. Dmitruk, and P. A. Lotito. Quadratic order conditions for bang-singular extremals. *Numer. Algebra Control Optim.*, 2(3):511–546, 2012.
- [9] B. Bonnard, J.-B. Caillaud, and E. Trélat. Second order optimality conditions in the smooth case and applications in optimal control. *ESAIM Control Optim. Calc. Var.*, 13(2):207–236 (electronic), 2007.
- [10] F. Chittaro and G. Stefani. Minimum-time strong optimality of a singular arc: the multi-input non involutive case. To appear in *ESAIM: COCV*, 2015.
- [11] F. C. Chittaro and G. Stefani. Singular extremals in multi-input time-optimal problems: a sufficient condition. *Control Cybernet.*, 39(4):1029–1068, 2010.
- [12] F. H. Clarke. On the inverse function theorem. *Pacific J. Math.*, 64(1):97–102, 1976.
- [13] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.
- [14] U. Felgenhauer. On stability of bang-bang type controls. *SIAM J. Control Optim.*, 41(6):1843–1867 (electronic), 2003.
- [15] J. Heinonen. *Lectures on Lipschitz analysis*, volume 100 of *Report. University of Jyväskylä Department of Mathematics and Statistics*. University of Jyväskylä, Jyväskylä, 2005.
- [16] M. R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.*, 1:525–581, 1951.
- [17] T. Iwaniec and G. Martin. *Geometric function theory and non-linear analysis*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2001.
- [18] J. M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.
- [19] H. Maurer and N. P. Osmolovskii. Second order sufficient conditions for time-optimal bang-bang control. *SIAM J. Control Optim.*, 42(6):2239–2263 (electronic), 2004.
- [20] L. Poggolini. On local state optimality of bang-bang extremal. *Rend. Semin. Mat. Univ. Politec. Torino*, 64(1):1–23, 2006.
- [21] L. Poggolini and M. Spadini. Strong local optimality for a bang-bang trajectory in a Mayer problem. *SIAM J. Control Optim.*, 49(1):140–161, 2011.
- [22] L. Poggolini and M. Spadini. Bang-bang trajectories with a double switching time in the minimum time problem. To appear in *ESAIM: COCV*, 2015.
- [23] L. Poggolini and G. Stefani. State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Systems Control Lett.*, 53(3-4):269–279, 2004.
- [24] L. Poggolini and G. Stefani. Bang-singular-bang extremals: sufficient optimality conditions. *J. Dyn. Control Syst.*, 17(4):469–514, 2011.

- [25] L. Poggiolini and G. Stefani. A case study in strong optimality and structural stability of bang-singular extremals. In *Geometric control theory and sub-Riemannian geometry*, volume 5 of *Springer INdAM Ser.*, pages 333–350. Springer, Cham, 2014.
- [26] G. Stefani. Strong optimality of singular trajectories. In *Geometric control and nonsmooth analysis*, volume 76 of *Ser. Adv. Math. Appl. Sci.*, pages 300–326. World Sci. Publ., Hackensack, NJ, 2008.
- [27] G. Stefani and P. Zezza. Time optimality of a bang-bang trajectory with Maple. In *Lagrangian and Hamiltonian methods for nonlinear control 2003*, pages 135–140. IFAC, Laxenburg, 2003.
- [28] H. Whitney. *Geometric integration theory*. Princeton University Press, Princeton, N. J., 1957.